AD A 295 704

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	4/13/95	Final 1 Feb 93 - 15 Jan 95

4. TITLE AND SUBTITLE

Computational Problems in Image Analysis, Multiple Object Recognition, and Speech Recognition

5. FUNDING NUMBERS

DAAH04-93-G-0046

6. AUTHOR(S)

Basilis Gidas

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Brown University
Providence, Rhode Island 02912

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

31078.3-MA-SDI

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

During the period of the grant, 2/1/93 - 1/15/95, we developed: (1) a Bayesian framework for object detection and tracking; the algorithm was successfully tested on real - data in the detection and tracking of vehicles on a highway: (2) a recognition algorithm based on stochastic hierarchical, context - free - grammars type, object representation; the study has required the development of feasible pruning techniques for dynamic programming; (3) a new acoustic model for speech recognition based on a wavelet representation of the acoustic signal, and nonparametric prediction techniques.

14. SUBJECT TERMS

object detection, tracking, recognition, dynamic programming, context - free - grammar

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FINAL REPORT

Summary of Research

Description:

Our research supported by the ARO/SDI grant has had several successes in its two principal lines of investigation: model-based *object detection/tracking/recognition*, and *speech recognition* via nonparametric statistical techniques. A major goal of our research has been the development of a sound and unified theoretical basis for the design of models and algorithms, and for overcoming the underlying massive computational and combinatorial problems. The modelling and algorithms have been strongly influenced by, and have been implemented on, real - world applications. The parallel study of object and speech recognition has benefited both areas. Our main projects and contributions to application and theory, may be divided into three groups:

1. Object Detection and Tracking: We have explored a statistical Bayesian framework for simultaneously describing and tracking objects, on the basis of image sequence frames. The framework has been successfully tested in a highway traffic scenario (See Images 1 and 2). It involves two major components: *object models*, and *spatial - temporal data models.*

Our object models and *deformable templates.* Vehicles, for example, are represented by *prototypes*, but their silhouettes on the 2-D image plane exhibit a great deal of variability depending on the object's distance and orientation relative to the camera. These variabilities are articulated via a *prior distribution* on the "shape space". The *spatial - temporal* data models and designed using three (or more) consecutive frames at a time. To deal with the variability of the observed grey - levels due to variations in heightening, contrast, texture, and other effects, we employe *nonparametric statistics* such as rank tests and the Kolmogorov - Simirnov statistic. In addition to the random variation of "shape" and image data, in the highway problem there is a third variability: the number of vehicles in a given frame is unknown, and it may vary from frame to frame. This is treated by using a Poisson type process.

2. Object Recognition: We have completed an Xlib - based, graphic interface computer program for recognizing 2-D objects in environments highly

degraded by noise, blur, clutter, and occlusion. The algorithm has been tested an a small database of 2-D industrial tools such as pliers, hammers, screws, etc.; the results have been encouraging.

The recognition framework emphases *object representation, data models*, and algorithmic issues. These are briefly as follows:

(1) The object representation is based on *Stochastic Hierarchical Models* (SHM) which are variants of Stochastic Context - free -Grammars (in the Chomsky hierarchy of grammars). Our SHM's have two levels of hierarchy and syntax. The first level (top level) views an object as a concatenation of its articulated joints and parts; it is represented by a directed graph structure—called the *membership graph.* Each node in the membership graph is associated with a "high-level" primitive (i.e. a component part) that may be common to several objects; the arcs of the membership graph correspond to syntactic constraints relating the various parts—the constraints are either topological (qualitative) or geometric. The second level of the hierarchy serves to represent the boundaries of the high-level primitives by a cascade of "lower-level" primitives or units starting with local edges ("edgelets") which are concatenated to give small line segments ("linelets") which, in turn, are concatenated to give more global boundaries or surfaces. The entire concatenation process is represented by a Markov process with "jumps", which allows one boundary segment to terminate and change ("jump") into another boundary segment. (2) The lower-level elementary units interact directly with basic local description of the grey-level image data. The local data description are properly designed nonparametric statistics, i.e. local functions of data that are invariant under changes in imaging conditions and degradations. (3) The combination of SHM with the data models leads to a formulation of the recognition problem as a global optimization problem which, in view of the recursive structure, lends itself to variations of *dynamic programming.* The dynamic programming process involves a large state space, and it requires the maintenance of a multitude of intermediate data structures. This prohibits the possibility of exact computations, and hence efficient pruning procedures are required. We have developed various optimal and sub-optimal heuristics for pruning, using a multiresolution analysis procedure. The overall recognition algorithm leads to a *simultaneous* interpretation at multiple levels ("low-level" primitives and "high-level" complex entities); *no primitive at any level is determined until the entire computation is completed.*

3. Speech Recognition: We have developed a new *acoustic model* for speech recognition alternative to that is the HMM approach; it explores three basic tools: A *wavelet* representation of the raw signal, and its induced "*waveletogram*"; nonparametric transformations of the waveletogram; and 1-D Markov Random Field (MRF) models (analogous to Markov models for phonemes in the HMM approach). Most speech recognition procedures (including HMM) assume that short time segments of the acoustic signal are *stationary* and *linear.* Hence, the signal is analyzed via Fourier Transform (FT), and linear models such as Linear Predictive Coding (LPC). These procedures are adequate in some parts of the signal (e.g. study states of vowels), but not in other parts: *Nonstationarities* in burst and transition regions (e.g. consonant to vowel) make the application of FT questionable; and *nonlinearities* contain important information that cannot be captured by LPC. The former of these difficulties (nonstationarity) is alleviated by using wavelets, while the later points to nonparametric statistics. The output of the nonparametric transformations may be viewed as a (compressed) process which is modeled by appropriate 1-D MRF's. Our procedure has also been applied to an important linguistic task: The classification of the six stop consonants |p, t, k, b, d, g| on the basis of CV (Consonant - Vowel) or VC syllables. Our procedure yields interesting 2-D clustering plots for vowels and consonants. We know of no other method in the interactive that gives such scatterplots for stop consonants.

LIST OF IMAGES

Single object segmentation and tracking (times 2, 4, 6 and 8)

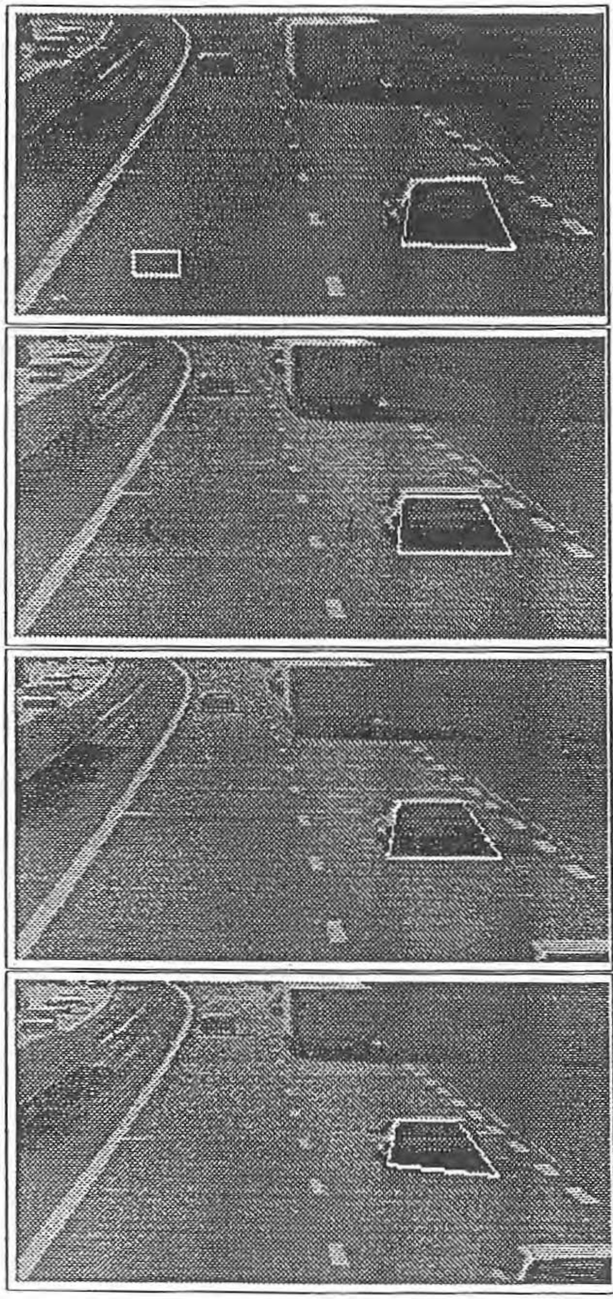

Figure 1

Three object segmentation and tracking (times 2, 4, 6 and 8)

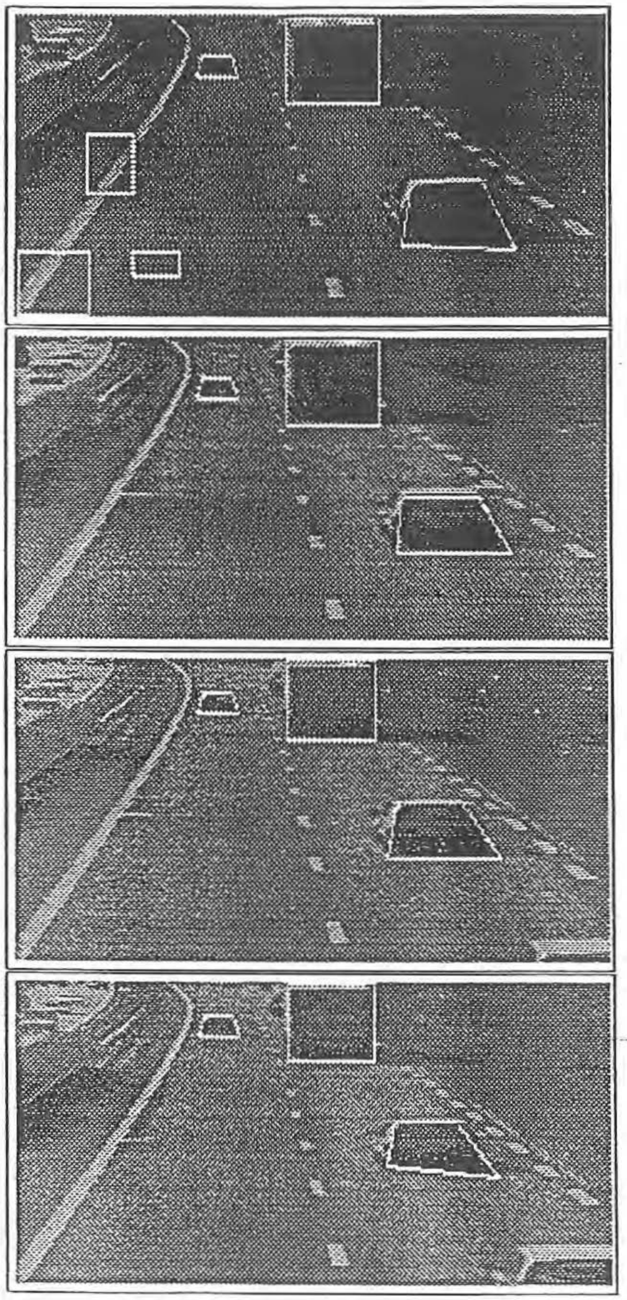

Figure 2

LIST OF PUBLICATIONS:

1. "A Variational Method for Estimating the Parameters of MRF from Complete or Incomplete Data" (with M. Almeida) Ann. Appl. Prob. 3 (1993) 103–136.
2. "Parameter Estimation for Gibbs Distributions from Fully Observed Data" in Markov Random Fields: Theory and Applications, Academic Press 1993, pp. 471–498
3. "Metropolis-type Monte Carlo Simulation Algorithms and Simulated Annealing", Topics in Contemporary Probability and Its Applications, CRC press 1995, 71 pages, ed.: J. L. Snell
4. "Motion Detection and Tracking Using Deformable Templates" (with P. Perez), Proceed. 1994 IEEE Intern. Conf. on Image Processing, Austin, Texas, Vol. II, pp 272-276.
5. Stop Consonants Discrimination and Clustering Using Nonlinear Transformations and Wavelets", (with A. Murua), Proceed on Image Models and their Speech Cousins, IMA, University of Minnesota, t appear 1995, eds: L. Shepp and S. Levinson.
6. "Classification and Clustering of Stop Consonants via Nonparametric Transformations" (with A. Murua), Proceed. 1995 IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing, Detroit, Michigan.
7. "Discussion of Analysis and Reconstruction of Medical Images Using Prior Information", V. Johnson et. al., Bayesian Statistics in Science and Technology: Case Studies, Springer-Verlag 1995, ed.: R. Kass
8. "Nonparametric Estimations for Linear Predictors from a Finite Data Set, and Consistency as the Sampling Period Tend to Zero" (with A. Murua), preprint 1995..

LIST OF PARTICIPANTS

1. Dr. Alejandro Murua, Visiting Assistant Professor
2. Antonio Zelic, Ph.D. candidate
3. Gillen Patric, Ph.D. candidate
4. Chris Robertson, Ph.D. candidate